The S2S Framework

A framework for integrating tools and data services for search interface development

Eric Rozell, Tetherless World Constellation

Part 1 Outline

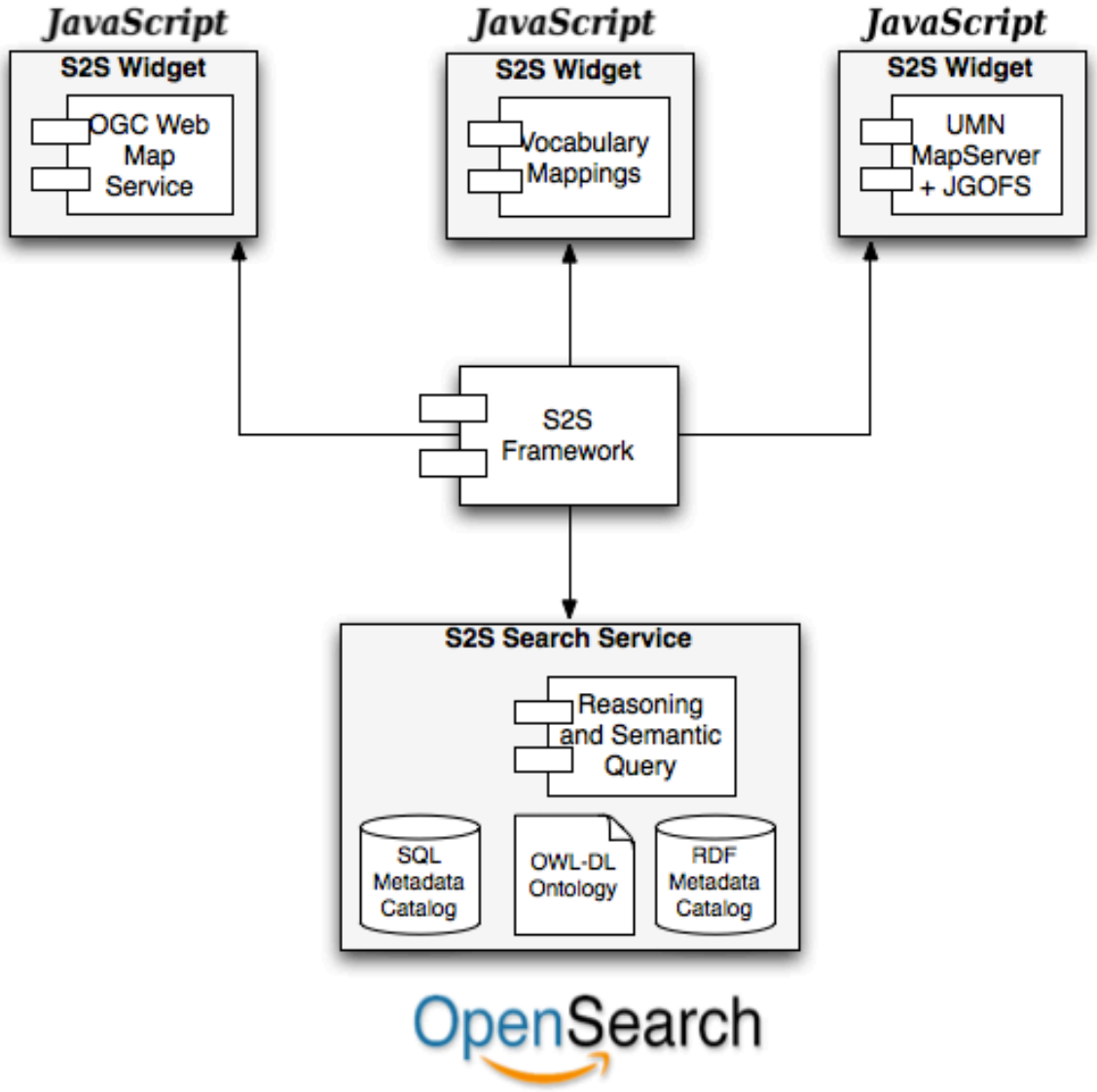
- Use Case
- Ontology
- Architecture
- OpenSearch in S2S
- Extensibility
- Demo
- Part 2

Use Case

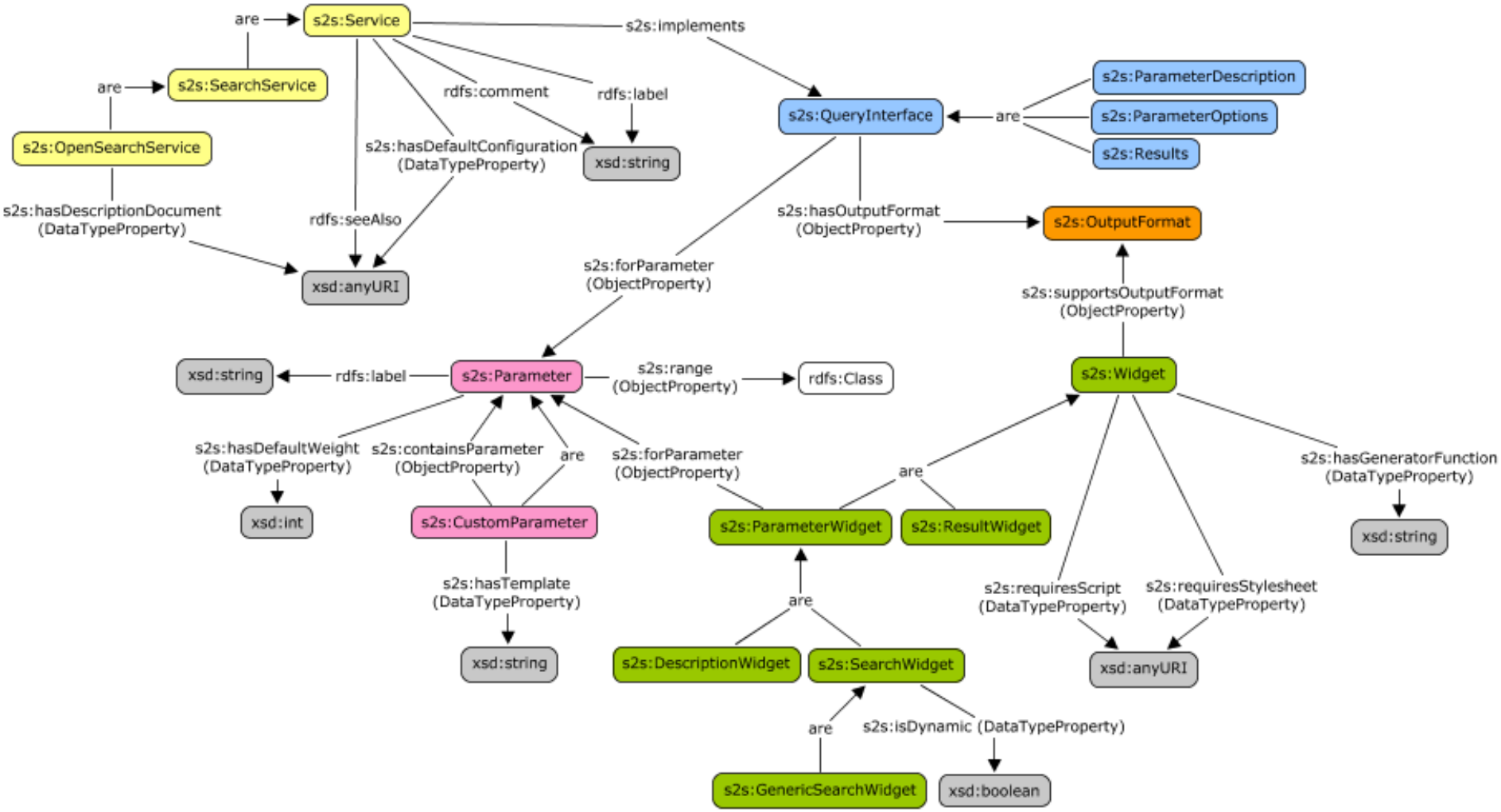
- BCO-DMO
  Biological and Chemical Oceanography Data Management Office
  - Build a faceted browsing interface to implement a variety of use cases, e.g.,
    - "Find datasets with containing [Parameter] from [Geographic Region]"
    - "Find datasets collected by [Person] funded by [award]"
  - Leverage existing technologies deployed by BCO-DMO

BCO-DMO
Biological and Chemical Oceanography Data Management Office

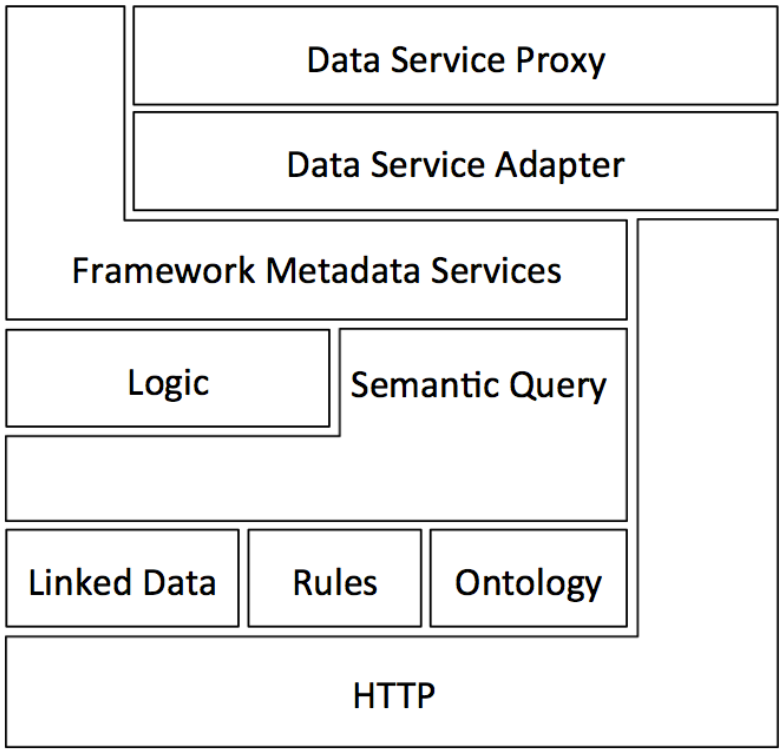
- Technologies (…or acronym soup)
  - JGOFS Data Server
  - SQL Metadata Catalog
  - UMN MapServer
  - OGC Web Map Service
  - OWL-DL Ontology
  - RDF variant of metadata
  - Vocabulary Mappings (R2R, SeaVoX)

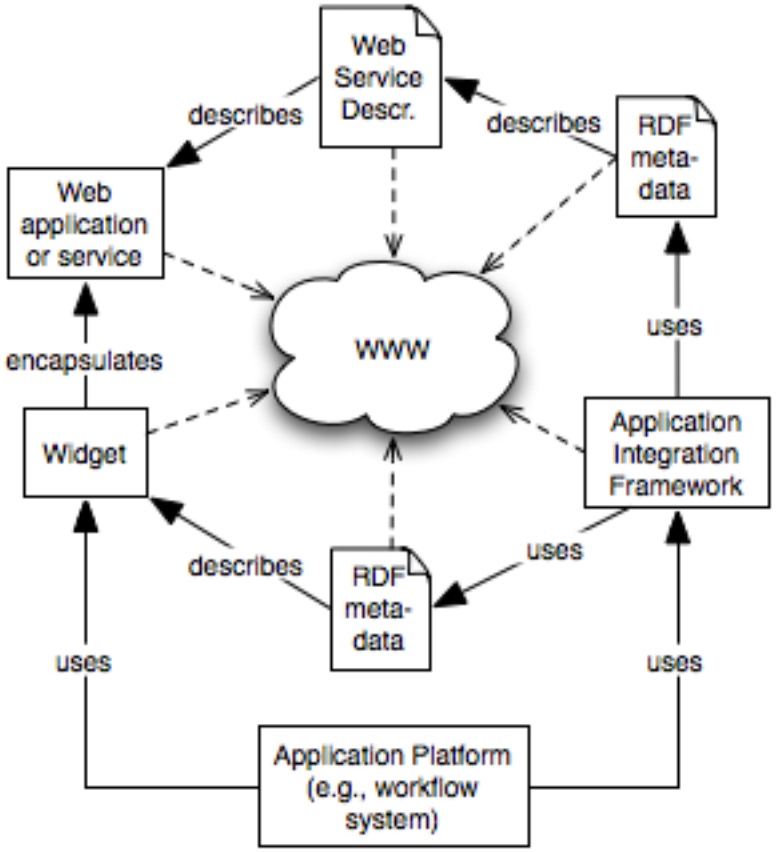

JavaScript
S2S Widget
OGC Web Map Service
JavaScript
S2S Widget
Vocabulary Mappings
JavaScript
S2S Widget
UMN MapServer + JGOFS
S2S Framework
S2S Search Service
Reasoning and Semantic Query
SQL Metadata Catalog
OWL-DL Ontology
RDF Metadata Catalog
OpenSearch

Ontology


are
s2s:Service
s2s:implements
are
s2s:SearchService
rdfs:comment
rdfs:label
s2s:QueryInterface
are
s2s:ParameterDescription
s2s:ParameterOptions
s2s:Results
s2s:OpenSearchService
s2s:hasDefaultConfiguration
(DataTypeProperty)
xsd:string
s2s:hasDescriptionDocument
(DataTypeProperty)
rdfs:seeAlso
s2s:hasOutputFormat
(ObjectProperty)
s2s:OutputFormat
xsd:anyURI
s2s:forParameter
(ObjectProperty)
s2s:supportsOutputFormat
(ObjectProperty)
xsd:string
rdfs:label
s2s:Parameter
s2s:range
(ObjectProperty)
rdfs:Class
s2s:Widget
s2s:hasDefaultWeight
(DataTypeProperty)
s2s:containsParameter
(ObjectProperty)
are
s2s:forParameter
(ObjectProperty)
are
s2s:hasGeneratorFunction
(DataTypeProperty)
xsd:int
s2s:CustomParameter
s2s:ParameterWidget
s2s:ResultWidget
xsd:string
s2s:hasTemplate
(DataTypeProperty)
are
s2s:requiresScript
(DataTypeProperty)
s2s:requiresStylesheet
(DataTypeProperty)
xsd:string
s2s:DescriptionWidget
s2s:SearchWidget
xsd:anyURI
are
s2s:isDynamic (DataTypeProperty)
s2s:GenericSearchWidget
xsd:boolean

Architecture





OpenSearch in S2S

- Connects to “rel” attributes and template parameters
- URIs must dereference to RDF descriptions (using S2S ontology!)

OpenSearch in S2S

- “rel” attributes contain s2s:QueryInterface instances, e.g.
  - http://a9.com/-/spec/opensearch/1.1/results
  - http://escience.rpi.edu/ontology/sesf/s2s/2/0/DataTableResults
- Template parameters contain s2s:Parameter instances, e.g.
  - http://a9.com/-/spec/opensearch/1.1/searchTerms
  - http://a9.com/-/opensearch/extensions/geo/1.0/box
  - http://a9.com/-/opensearch/extensions/time/1.0/start

Extensibility

- Framework designed to be generic
- Three primary extensibility axes
  - Widget Development (e.g., new JavaScript)
  - Service Development (e.g, new OpenSearch)
  - Adapter Development (e.g., add SAWSDL support)

Demo

- http://bit.ly/twiogdc
- http://bit.ly/bcodmo-demo

Linked Open Research Data for Earth and Space Science Informatics

Progress for the ESIP FUNding Friday Project

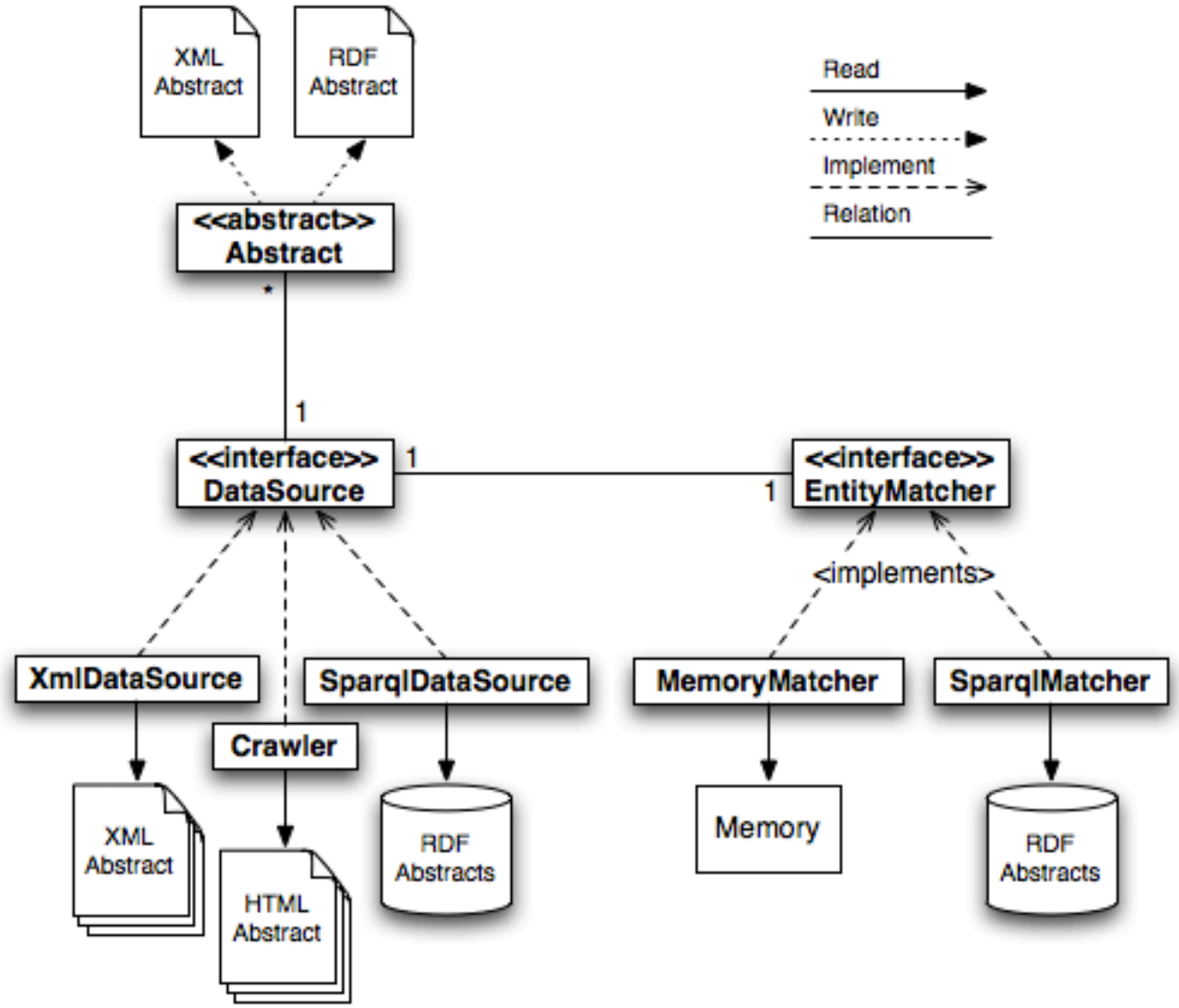
Part 2 Outline

- Goals
- Conversion Architecture
- Linked Data Publication
- RDF Vocabulary Design
- Named Entity Identification
- Linking ESIP Membership
- Demo

Goals

- Improve Access to AGU Abstracts
  - Provide various data formats
  - 5-Star Linked Data Publishing
- Identify Topic Trends in ESSI
  - Identify technology “mentions”
  - Link technology “mentions” to external vocabularies
- Enhance Expert/Collaboration Discovery
  - Link authors to members of external communities
  - Visualize impact/participation in informatics topics

Conversion Architecture



Linked Data Publication

★	make your stuff available on the web (whatever format)
★★	make it available as structured data (e.g. excel instead of image scan of a table)
★★★	non-proprietary format (e.g. csv instead of excel)
★★★★	use URLs to identify things, so that people can point at your stuff
★★★★★	link your data to other people's data to provide context

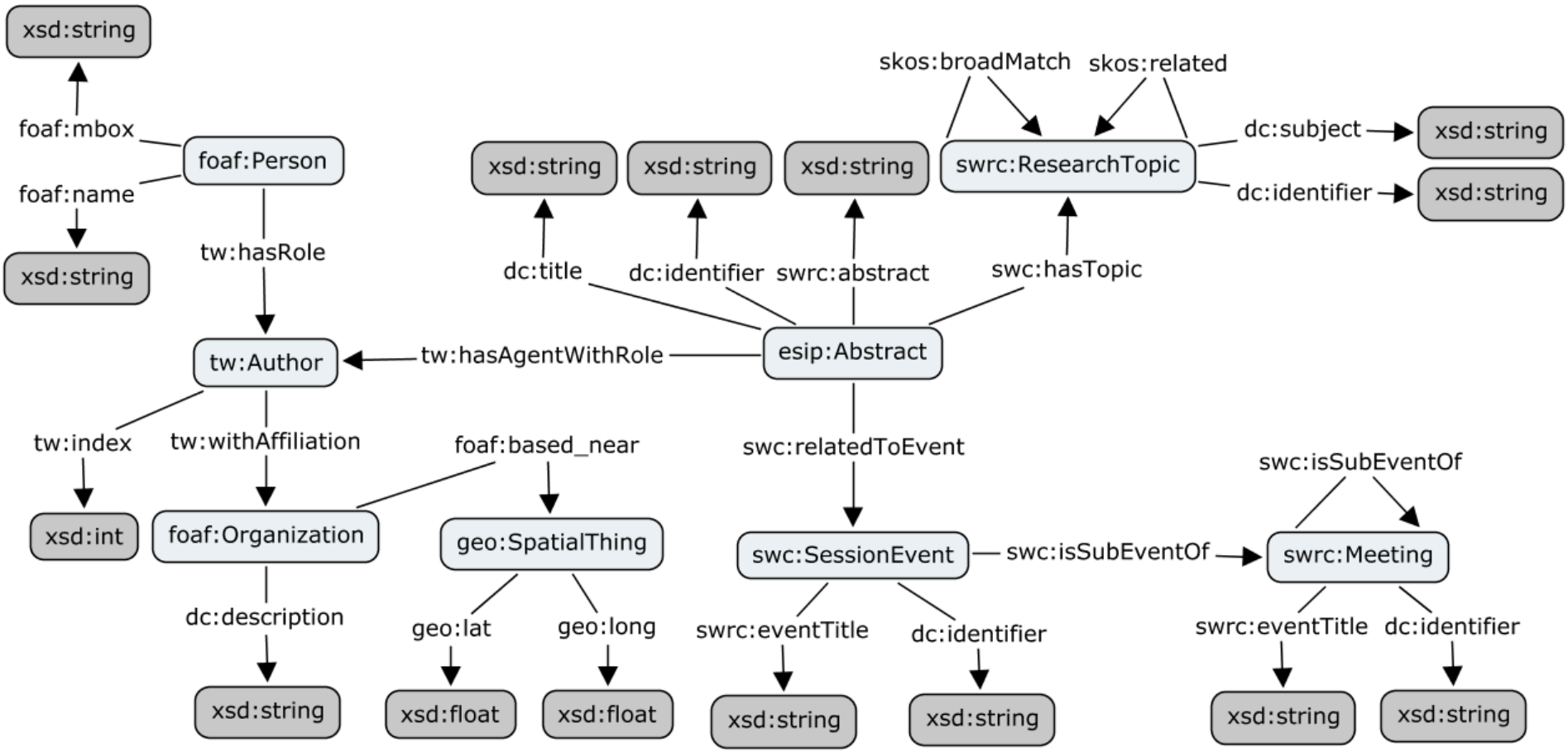
Linked Data Publication

- Simple Publishing Kit for Linked Open Data (LODSPeaKr)
  - Easy to deploy from existing SPARQL endpoint
  - Uses query “models” to collect data for a class of RDF instances
  - Uses “views” based on Haanga templates to visualize RDF as XHTML, RDF/XML, etc.
- Our Integration
  - Point LODSPeaKr at our endpoint
  - Write models and views for HTML pages
  - Using default behavior for RDF responses

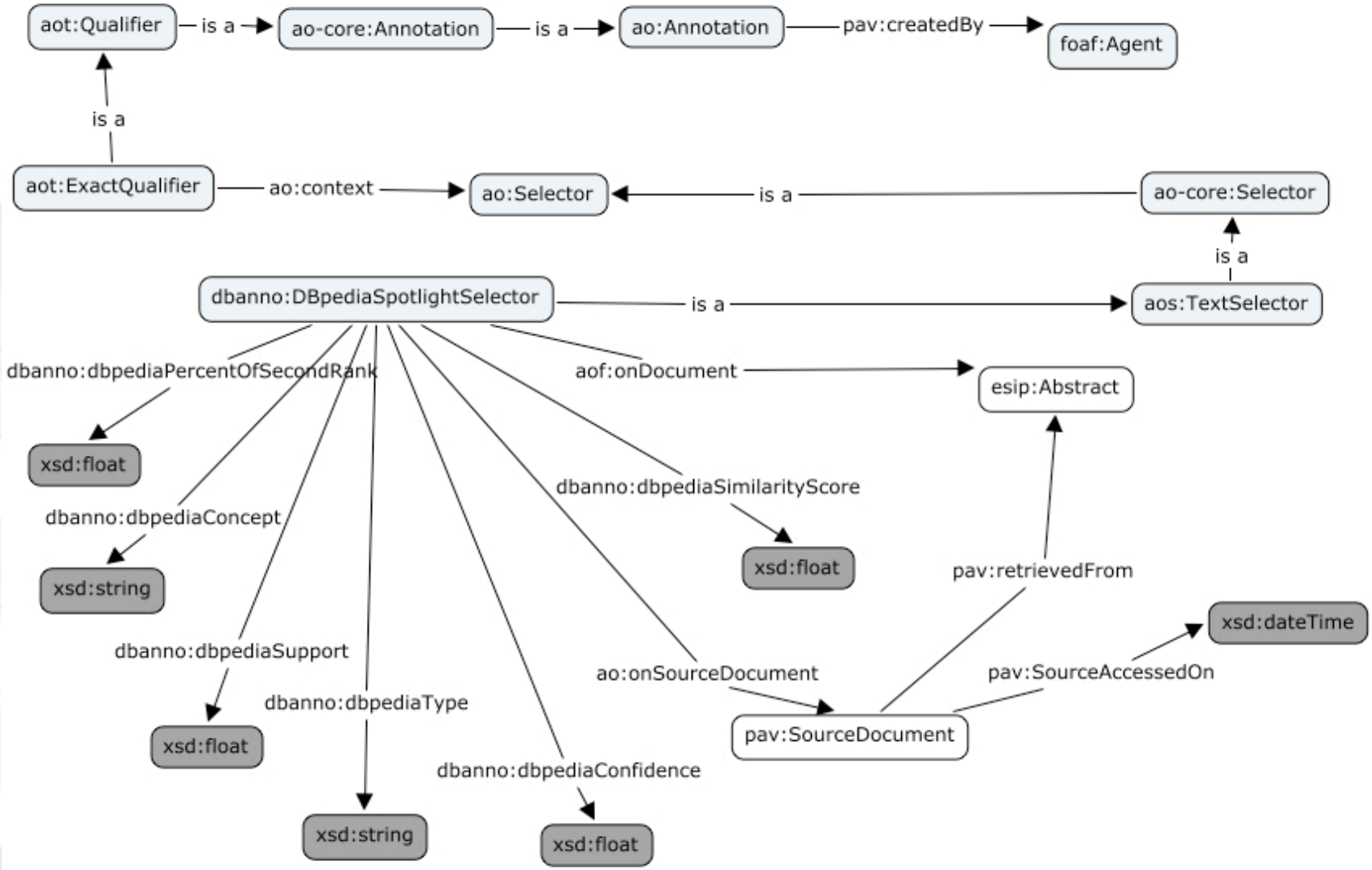
RDF Vocabulary Design

- foaf – Friend of a Friend
- dc – Dublin Core
- swrc – Semantic Web for Research Communities
- swc – Semantic Web Conference
- tw – Tetherless World
- geo – WGS84 Lat/Long
- skos – Simple Knowledge Organization System
- xsd – XML Schema
- ao – Annotation Ontology

RDF Vocabulary Design



RDF Vocabulary Design



Named Entity Identification

- DBPedia Spotlight
  - Combines surface form spotter and disambiguation algorithm
  - Computes confidence in disambiguations
  - Offered as Java package and Web service
- Our Integration
  - Using Web service and confidence threshold
  - Convert response into RDF
  - Store RDF with our existing abstract data

Linking ESIP Membership

- From ESIP meeting attendance logs
- “EntityMatcher” infrastructure to match ESIP members to ESSI authors
- Currently using simple rules, i.e., “if email matches, then person matches”
  - Looking to more complex rules, e.g., “if name and affiliation is similar, and belongs to similar co-author clique”
- owl:sameAs statements

Demo

- http://bit.ly/essi-lod-demo